

Title:

The COVID-19 mortality effects of underlying health conditions in India: a modelling study

Authors:

Corresponding Author:

Paul Novosad, Ph.D. (0000-0001-8331-6630)

Assistant Professor of Economics

Department of Economics

Dartmouth College

311 Rockefeller Hall

Hanover, NH 03755

+1 857 244 1383

paul.novosad@dartmouth.edu

Radhika Jain, Sc.D.

Postdoctoral Fellow

Shorenstein APARC

Stanford University

Encina Hall E301

Stanford, CA 94305-6055

rjain19@stanford.edu

Alison Campion, B.A.

Research Associate

Development Data Lab

1807 California St. NW, #106

Washington, DC 20009

acampion@alumni.princeton.edu

Sam Asher, Ph.D.

Assistant Professor of Economics

Department of Economics

Johns Hopkins University School of Advanced International Studies

Wyman Park Building 5th Floor, 3100 Wyman Park Drive

Baltimore, MD 21211

sasher2@sais.edu

Word Count: 3527

ABSTRACT

Objective: To model how known COVID-19 comorbidities will affect mortality rates and the age distribution of mortality in a large lower middle income country (India), as compared with a high income country (England), and to identify which health conditions drive any differences.

Design: Modelling study.

Setting: England and India.

Participants: 1,375,548 respondents aged 18 to 99 to the District Level Household Survey-4 and Annual Health Survey in India. Additional information on health condition prevalence on individuals aged 18 to 99 was obtained from the Health Survey for England and the Global Burden of Diseases, Risk Factors, and Injuries Studies (GBD).

Main outcome measures: The primary outcome was the proportional increase in age-specific mortality in each country due to the prevalence of each COVID-19 mortality risk factor (diabetes, hypertension, obesity, chronic heart disease, respiratory illness, kidney disease, liver disease, and cancer, among others). The combined change in overall mortality and the share of deaths under 60 from the combination of risk factors was estimated in each country.

Results: Relative to England, Indians have higher rates of diabetes (10.6% vs. 8.5%), chronic respiratory disease (4.8% vs. 2.5%), and kidney disease (9.7% vs. 5.6%), and lower rates of obesity (4.4% vs. 27.9%), chronic heart disease (4.4% vs. 5.9%), and cancer (0.3% vs. 2.8%). Population COVID-19 mortality in India relative to England is most increased by diabetes (+5.4%) and chronic respiratory disease (+2.3%), and most reduced by obesity (-9.7%), cancer (-3.2%), and chronic heart disease (-1.9%). Overall, comorbidities lower mortality in India relative to England by 9.7%. Accounting for demographics and population health explains a third of the difference in share of deaths under age 60 between the two countries.

Conclusions: Known COVID-19 health risk factors are not expected to have a large effect on aggregate mortality or its age distribution in India relative to England. The high share of COVID-19 deaths from people under 60 in low- and middle-income countries (LMICs) remains unexplained. Understanding mortality risk associated with health conditions prevalent in LMICs, such as malnutrition and HIV/AIDS, is essential for understanding differential mortality.

Keywords: COVID-19, India, low- and middle-income countries, comorbidity

SUMMARY BOX

What is already known on this topic

COVID-19 infections in low- and middle-income countries (LMICs) are rising rapidly, with the burden of mortality concentrated at much younger ages than in rich countries.

A range of pre-existing health conditions can increase the severity of COVID-19 infections.

It is feared that poor population health may worsen the severity of the pandemic in LMICs.

What this study adds

The COVID-19 comorbidities that have been studied to date may have only a very small effect on aggregate mortality in India relative to England and do not shift the mortality burden toward lower ages at all.

India's younger demographics can explain only a third of the substantial difference in the share of deaths under age 60 between India and England.

However, mortality risk associated with health conditions prevalent in LMICs, such as malnutrition and HIV/AIDS, is unknown and research on this topic is urgently needed.

INTRODUCTION

As governments around the world ease social distancing measures imposed to limit the transmission of COVID-19, the number of global cases is rising. A growing share of cases is now coming from low- and middle-income countries (LMICs) in Asia, Africa, and the Americas that were largely spared in the initial stages of the pandemic.¹ Because the severity of infection is substantially increasing with age, forecasts have projected much lower aggregate mortality rates in LMICs than in wealthier countries.²⁻⁴

However, reported fatality numbers from LMICs to date have suggested a much greater share of COVID-19 deaths among the young. 30.5% of deaths in Brazil and 50% of deaths in India have occurred in those under age 60; 27% of deaths in Mexico have occurred in those under 50.⁵⁻⁷ In contrast, individuals under 65 have accounted for only 5–13% of deaths in 10 European countries and Canada and 8–24% in US states.⁸ It is not presently known whether the different age pattern of deaths in LMICs is driven by erroneous reporting, differences in infection patterns, younger populations, or worse underlying population health.

Many modelling studies have presumed that worse population health in poor countries will lead to excess mortality, or else ignored differential population health as a factor entirely.^{4,9,10} To date there has been limited analysis of the prevalence in LMICs of the specific conditions associated with increased COVID-19 severity, such as diabetes, obesity, cardiovascular disease, hypertension, and chronic kidney disease, nor of how they change the expected level and age distribution of mortality.^{4,11-14} Some studies have adjusted mortality estimates for population comorbidities by treating all comorbidities as equivalent or by multiplying the mortality rate by a fixed amount to adjust for population health.¹⁵⁻¹⁷ One study combined condition-specific prevalence and hazard ratios from a sample of hospitalizations, but excluded obesity and uncontrolled diabetes, and did not examine mortality or the age distribution of mortality as outcomes.¹⁸

Using England as a benchmark, this study examines how comorbidities understood to increase COVID-19 mortality are likely to affect COVID-19 mortality rates in aggregate and across the age distribution in India, identifying the specific risk factors with the largest mortality effects. We further study the extent to which accounting for differences in demographics and underlying health conditions can explain the increased share of deaths among the young in India relative to England.

Our analysis focuses on India and on the COVID-19 risk factors that are currently documented. India presently has the third highest number of cumulative COVID-19 infections in the world and the highest growth rate in infections of any major country, making it an essential population to study.¹ The methodology is readily adjusted to account for new risk factors or data from other countries and may be useful for modelling the epidemic in a range of LMICs.

METHODS

Our approach requires three types of data: (i) the relative risk of COVID-19 mortality associated with gender, age, and each health condition; (ii) the age-specific prevalence of each health condition in England and India; and (iii) the age and gender distributions for the two countries.

Estimates of Relative Risk of COVID-19 Mortality from Comorbidities

We obtained estimates of COVID-19 mortality risk for a wide range of comorbidities from the OpenSAFELY study, a closed cohort study of 17,425,445 adults from England.¹⁹ This was the largest analysis of comorbidities associated with COVID-19 mortality to date and one of the few studies that estimates risk factors in a multivariate model adjusting for age, sex, and other health conditions. This

adjustment is important because many COVID-19 comorbidities are increasing in age and their hazard ratios are thus biased upward in analyses not adjusting for age.

The OpenSAFELY study enrolled all individuals registered with a general practice within The Phoenix Partnership system on 1st February 2020, were 18 years or older upon enrolment, had at least one year of prior medical history within the system, and had recorded age and sex. The underlying dataset represents 40% of the population of England and the prevalence of health conditions in the study cohort is similar to estimates of population prevalence in England (appendix p 4). Patients were followed through 25th April. The outcome was in-hospital death among people with confirmed COVID-19 infections. Hazard Ratios (HRs) for mortality from a Cox proportional hazards model were estimated for a comprehensive list of risk factors described in other studies, adjusted for sex, age, and all other risk factors. As patient-level data from OpenSAFELY are not publicly available, we extracted HRs from the paper reporting results of the analysis.¹⁹

Ideally, HRs would measure mortality risk conditional upon infection, rather than on clinic attendance (this paper) or hospitalization (as in prior work).^{18,20} The HRs in this study reflect combined mortality and infection risk; the analysis assumes that pre-existing health conditions are not significant predictors of infection. However, condition HRs measuring mortality risk conditional upon hospitalization are similar to those used here.¹¹

Demographics and Risk Factor Prevalence in India and England

Age distributions and age-specific sex ratios for India and England were obtained from official censuses.

We obtained data on age-specific prevalence of health risk factors for India and England from multiple sources, prioritizing biomarker data where available, and matching definitions as closely as possible to the conditions for which HRs are available. We restrict samples to ages 18–99 for consistency with the HRs.

For India, we used biomarker data from two public population health surveys for obesity, diabetes, and hypertension. The fourth round of the Indian District Level Household Survey (DLHS-4) and the second round of the Annual Health Survey (AHS) were conducted between 2012 and 2014, jointly cover 94% of the Indian population, and provide the most recent nationwide direct measures of height, weight, fasting plasma glucose (FPG) and blood pressure (BP) for adults of all ages in India. Details of dataset construction are provided in the Appendix (p 1). For England, age-specific prevalences of obesity, hypertension, and diabetes were obtained from the nationally representative 2018 Health Survey for England, which collected symptoms and medical diagnoses for a range of illnesses, as well as direct measures of height, weight, blood pressure, and glycated haemoglobin (HbA1c).²¹

BMI was classified into no evidence of obesity (<30kg/m²), obese class one or two (30–39.9kg/m²), and obese class three (40+kg/m²). Hypertension was defined as systolic BP ≥140 mm Hg or diastolic BP ≥90 mm Hg (uncontrolled) or a medical diagnosis of hypertension with BP below the thresholds (controlled). Controlled and uncontrolled hypertension prevalence were reported separately but combined in the risk estimation for consistency with OpenSAFELY. OpenSAFELY classified controlled diabetes as HbA1c > 51 mmols/mol and HbA1c < 58mmols/mol, and uncontrolled diabetes as HbA1c ≥ 58 mmols/mol. Corresponding thresholds for the one-time FPG measures in the Indian dataset are not well defined. In England, prevalence was reported based on a threshold of HbA1c ≥ 48mmol/mol (6.5%). Therefore, we followed the standard screening and diagnosis thresholds recommended by the WHO and International Diabetes Federation and defined uncontrolled diabetes in India as a plasma glucose reading ≥126mg/dL [7.0mmol/L] if fasting or ≥200 mg/dL [11.1 mmol/L] if not fasting. We used the corresponding recommended threshold of HbA1c ≥48mmol/mol (6.5%) for uncontrolled diabetes in England.²² In both

countries, we classified individuals with biomarkers below the thresholds but with a diagnosis of diabetes as having controlled diabetes.

Age-specific prevalence for asthma, chronic heart disease, kidney disease, stroke, dementia, haematological malignancies, and all other cancers were drawn from the Global Burden of Diseases, Risk Factors, and Injuries Studies (GBD) for India and England.²³ OpenSAFELY reports separate HRs for cancers diagnosed <1 year ago, 1–4.9 years ago, and ≥5 years ago; because the year of diagnosis is unavailable in GBD, we used a single classification for each class of cancers and the HR for diagnosis <1 year ago. For chronic respiratory disease, we used COPD prevalence from the GBD for India and modelled COPD prevalence from the Clinical Practice Research Datalink cohort database for England.²⁴ GBD prevalence of Parkinson's disease, epilepsy, multiple sclerosis, and motor neuron disease were combined and classified as neurological disorders.

The following risk factors were not available for India and were excluded from the analysis for both England and India for comparability: fibrosing lung disease, bronchiectasis or cystic fibrosis, lupus, asthma with no recent OCS use, cancers diagnosed more than a year ago, organ transplant, and spleen disease. Given that the relationship between smoking and COVID-19 mortality remains under debate, we excluded it from the analysis.²⁵ We also excluded ethnicity and socioeconomic status, which cannot be measured comparably across England and India and are unlikely to have similar relative risk in the two countries.

Estimating the Contribution of Health Conditions to Population COVID-19 Mortality Risk

The OpenSAFELY study reports HRs for each age group, sex, and health condition with females 50–59 years with no conditions as the reference group.¹⁹ We transform the HR for each health condition c into a relative risk (RR_c) assuming a population mortality rate r of 1%:

$$RR_c = (1 - e^{HR_c \ln(1-r)})/r$$

To obtain continuous relative risk for age, we used a polynomial interpolation for the log HR at each age, renormalizing with age 50 as the reference group (appendix p 5).

The increase in population mortality risk from a given health condition is increasing in the condition's relative risk for COVID-19 mortality and its prevalence at each age. We defined the age- and condition-specific population relative risk $PRR_{a,c}$ of condition c at age a as:

$$PRR_{a,c} = RR_c * PREV_{a,c} + (1 - PREV_{a,c})$$

$PREV_{a,c}$ is the prevalence of condition c at age a . $PRR_{a,c}$ describes the proportional increase in mortality at age a driven by health condition c .

We combined PRRs to obtain an age-specific population relative risk of mortality arising from the combined prevalence of all of the health conditions :

$$PRR_a = \prod_{c \in \mathcal{C}} PRR_{a,c}$$

PRR_a isolates the expected mortality difference at each age between India and England that is driven by the combined prevalence of all of the health conditions studied. This approach implicitly assumes that the health conditions are uncorrelated with each other. Without microdata on the full set of health conditions, this assumption is unavoidable, but will bias the England vs. India comparison only if the correlation of

health conditions is substantially different in the two countries. We explore the possible extent of this bias in the appendix. By using age-specific prevalence, our analysis fully accounts for the substantial correlations between age and health conditions.

We next calculated the increase in population mortality from each health condition across all ages, taking into account the age-specific prevalence of each health condition, its relative risk, and the population share at each age. The condition-specific population relative risk of each health condition across the full population (PRR_c) is given by:

$$PRR_c = \sum_{a \in [18,99]} (PRR_{a,c} * POPSHARE_a)$$

PRR_c is greater when the relative risk of condition c is higher, and when its prevalence is higher at ages with higher population and greater mortality risk. The combined effect on population mortality of all of the health conditions is given by the product of each condition-specific PRR_c .

Finally, we aggregated the population relative risks across health conditions in order to model the age distribution of deaths in each country. The number of deaths at each age N_a is the product of the mortality rate of the reference group (50-year-old women with no other risk factors), the population at age a , the age-specific population relative risk of the full set of health conditions, the PRR of gender, and the direct relative risk of COVID-19 mortality for an individual at age a (RR_a):

$$N_a = r * POP_a * PRR_a * PRR_{a,male} * RR_a$$

We plotted the age distribution of deaths as shares of all deaths rather than in levels, eliminating the need to assume a reference group mortality rate. We summarized the shape of the distribution by reporting the share of expected deaths in each country that are under the age of 60. We present results from three models: (i) England's demographics and health distribution; (ii) India's demographics and health distribution; and (iii) India's demographics but England's age-specific prevalence of health risk factors. The third model allowed us to examine the mortality shift that comes from differences in population health alone.

Role of the Funding Source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing. The corresponding author had full access to the data in the study and had final responsibility for the decision to submit for publication.

Patient and Public Involvement

Because this study uses existing epidemiological data, it was not appropriate to involve patients or the public in the research.

RESULTS

Prevalence of COVID-19 risk factors in India and England

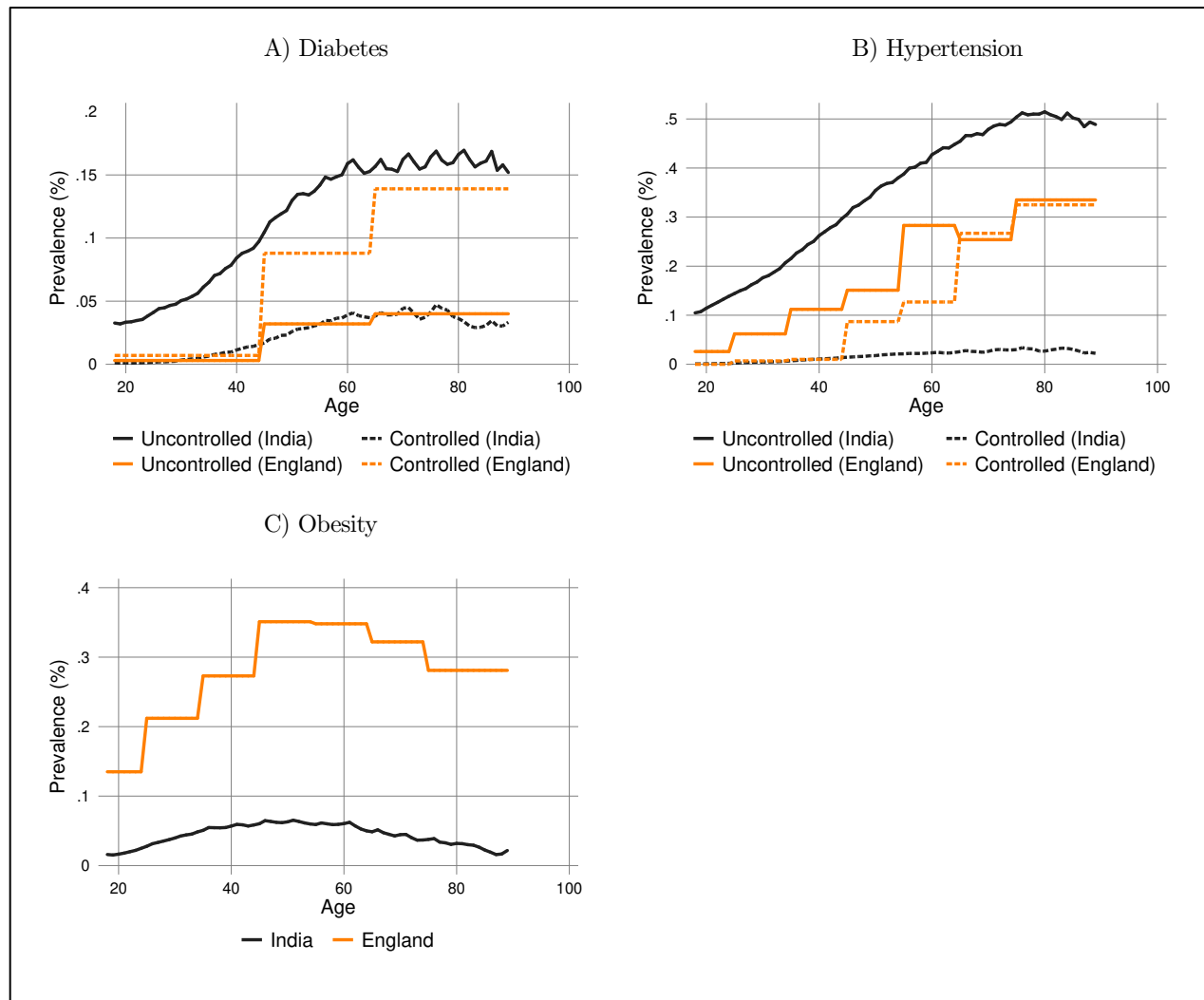
Demographic characteristics and overall prevalence of risk factors are substantially different in India relative to England (table 1). 83.8% of Indian adults are below the age of 60, compared to 69.9% of English adults. Indians have substantially lower rates of obesity and cancer (4.4% and 0.3% in India compared with 27.9% and 2.8% in England), but higher rates of uncontrolled diabetes, kidney disease, and chronic liver disease (8.9%, 9.7%, and 5.3% in India and 2.1%, 5.6%, and 2.6% in England).

Table 1
Prevalence of COVID-19 risk factors in India and England

	Prevalence (%)	
	India	England
Age 18-39	50.2	36.6
Age 40-49	19.2	16.3
Age 50-59	14.3	17.0
Age 60-69	10.3	13.3
Age 70-79	4.6	10.4
Age 80-99	1.5	6.3
Male	47.1	48.9
Diabetes (Controlled)	1.7	6.4
Diabetes (Uncontrolled)	8.9	2.1
Hypertension	28.2	28.0
Obese (class I & II)	4.0	24.8
Obese (class III)	0.4	3.1
Chronic Heart Disease	4.4	5.9
Chronic Respiratory Disease	4.8	2.5
Asthma	2.5	9.2
Kidney Disease	9.7	5.6
Chronic Liver Disease	5.3	2.6
Haematological Cancer	0.0	0.2
Non-haematological Cancer	0.3	2.6
Stroke, Dementia	1.3	1.5
Other Neurological Condition	0.0	0.1
Psoriasis, Rheumatoid	1.0	2.4
Other Immunosuppressive Conditions	0.1	0.1

We show age-specific prevalence differences between India and England for the conditions for which we have biomarkers in India and are more precisely estimated (figure 1), as well as age-specific prevalence of all conditions for both countries (appendix p 3). Overall rates of diabetes are higher in India at all ages, but diabetes in India is overwhelmingly uncontrolled, while three quarters of diabetes is controlled in England. Hypertension (the sum of controlled and uncontrolled) is higher in India at young ages (31.3% for ages 40–49 in India and 18.3% in England) but lower at higher ages (52.3% at ages 70–79 in India and 61.3% in England) and is overwhelmingly uncontrolled. Conversely, obesity rates are higher at all ages in England.

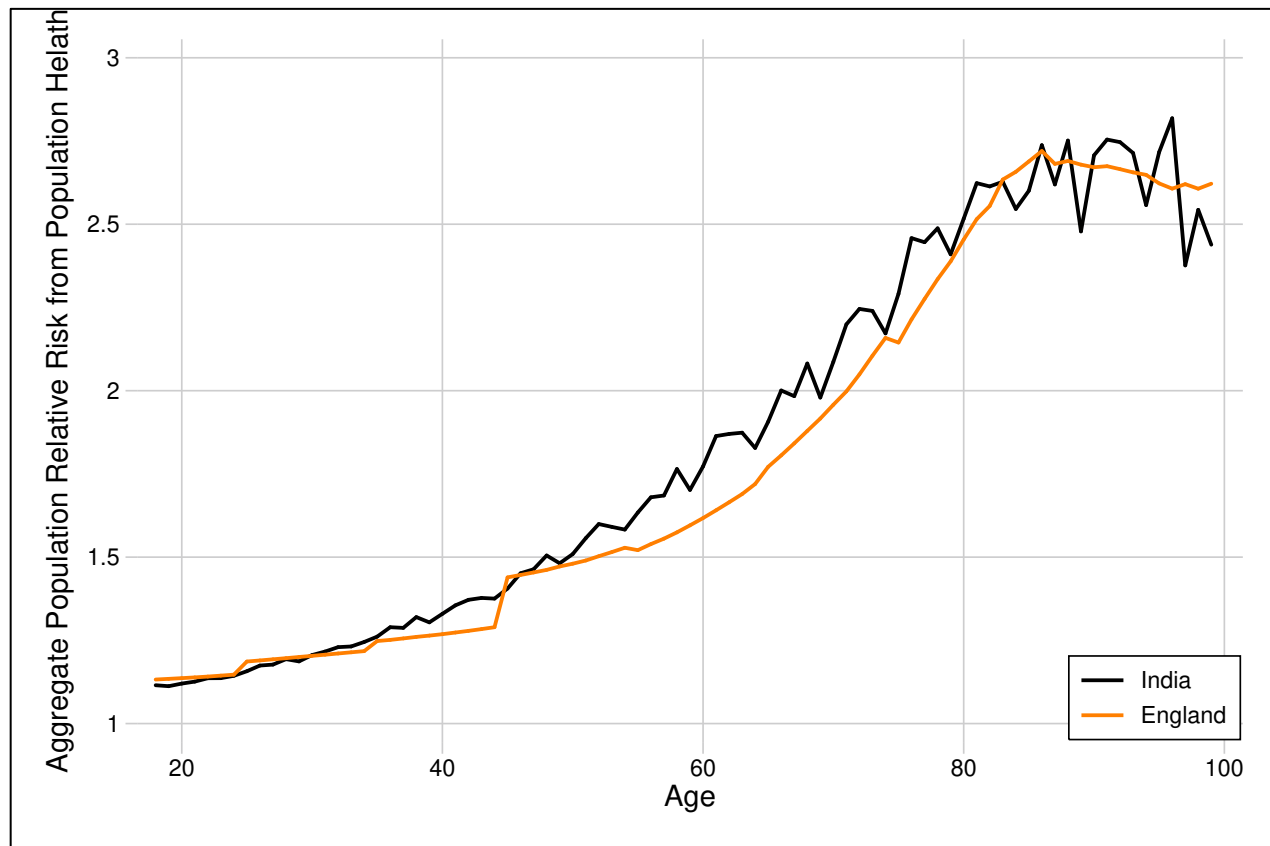
Figure 1
Prevalence of diabetes, hypertension, and obesity in India and England



Relative risk of COVID-19 mortality from combined risk factors in India and England

The age-specific population relative risk across all health conditions combined (PRR_a) is higher in India than in England at nearly all ages, but the difference in PRR_a between the two countries is below 15% at every age (figure 2). Modelled age-specific mortality rates in India are highest relative to England between ages 40 and 80.

Figure 2
Age-specific population relative risk from combined health conditions (PRR_a)



The PRR_a reflects the age-specific prevalence and associated risk of each health condition. Taking risk, prevalence, and population at every age into account provides the full population relative risk of each health condition (PRR_c) — the proportional increase in population mortality across all ages driven by each health condition (table 2).

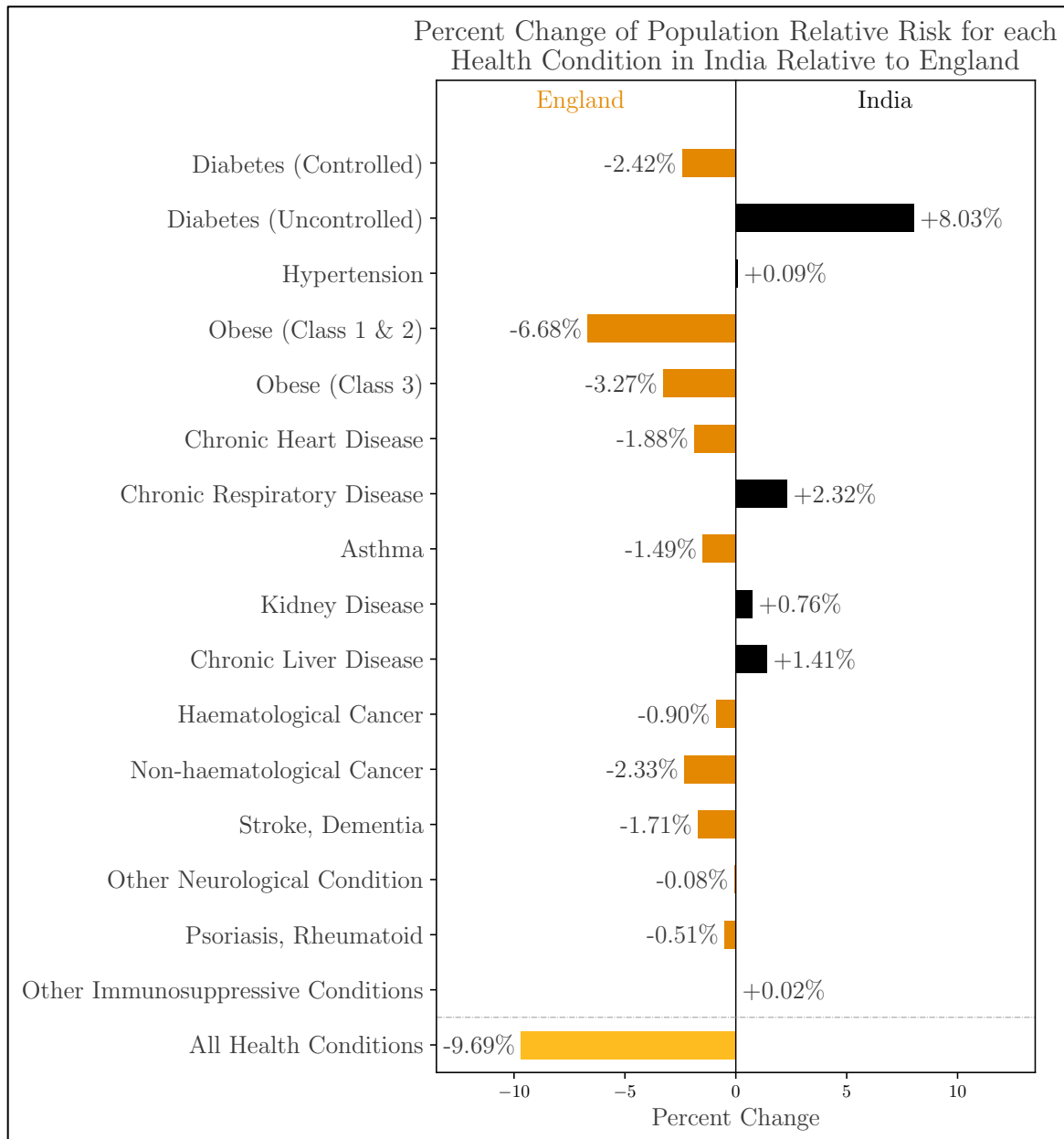
Table 2
Population relative risk of COVID-19 mortality from each health condition

	Individual Relative Risk	Population Relative Risk	
		India	England
Diabetes (Controlled)	1.50	1.007	1.032
Diabetes (Uncontrolled)	2.34	1.111	1.028
Hypertension	0.95	0.987	0.986
Obese (class I & II)	1.35	1.014	1.086
Obese (class III)	2.26	1.005	1.039
Chronic Heart Disease	1.27	1.013	1.033
Chronic Respiratory Disease	1.77	1.043	1.019
Asthma	1.25	1.006	1.022
Kidney Disease	1.71	1.086	1.078
Chronic Liver Disease	1.61	1.034	1.020
Haematological Cancer	3.48	1.001	1.010
Non-haematological Cancer	1.56	1.002	1.026
Stroke, Dementia	1.78	1.011	1.028
Other Neurological Conditions	2.44	1.002	1.002
Psoriasis, Rheumatoid	1.23	1.003	1.008
Other Immunosuppressive Conditions	1.68	1.001	1.002

Uncontrolled diabetes, which is associated with substantial mortality risk (RR_c 2.34), increases total mortality by 11% in India (PRR_c 1.11), but only 3% in England (PRR_c 1.03), reflecting its significantly higher prevalence in India at all ages. In contrast, controlled diabetes, more common in England than India, raises mortality by 3% in England and only 1% in India (PRR_c 1.03 vs 1.01). In addition to uncontrolled diabetes, the health conditions causing the largest increases in mortality in India are kidney disease ($PRR_c = 1.09$), and chronic respiratory disease ($PRR_c = 1.04$). In England, the most consequential health conditions are obesity (combined PRR_c across all obesity classes = 1.13), and kidney disease ($PRR_c = 1.08$).

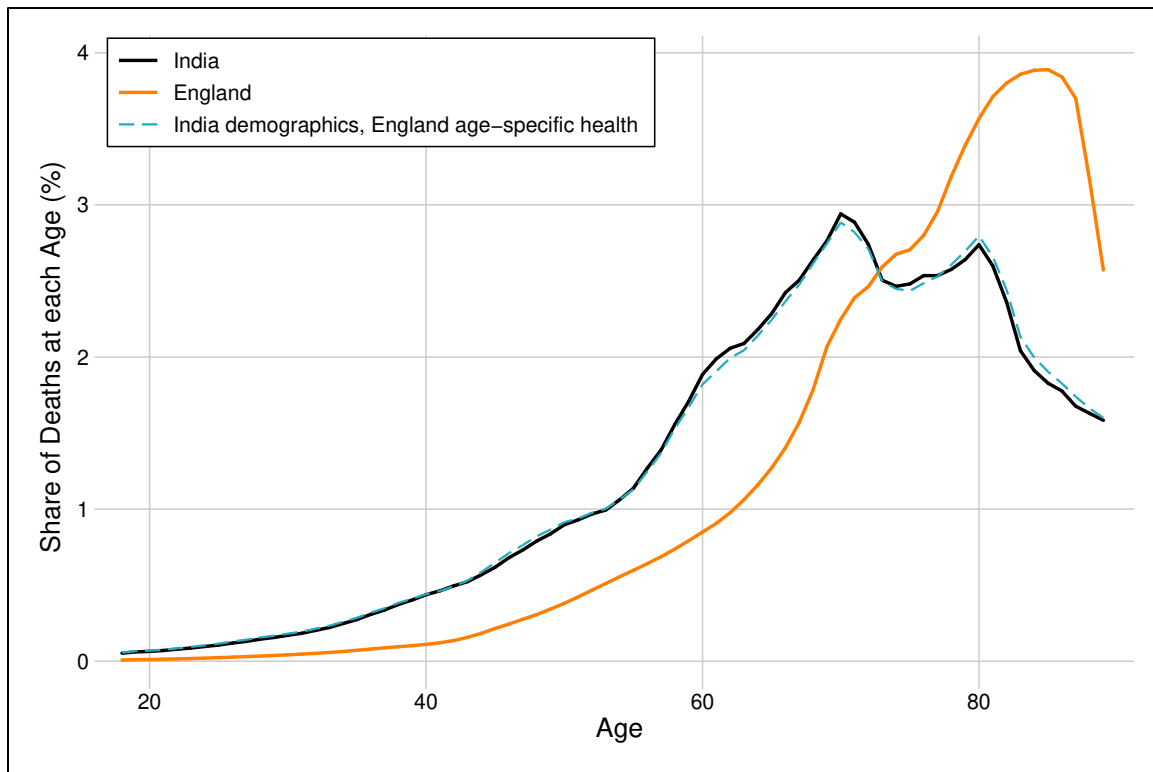
Comparing the percentage difference between the PRR_c of each health condition between India and England (figure 3), the condition with the largest differential impact on mortality between the two countries is uncontrolled diabetes, which increases population mortality by 8.03% in India relative to England. Mortality in India relative to England is also increased by chronic respiratory disease (+2.3%) and chronic liver disease (+1.4%), but decreased by the differential prevalence of obesity (combined -9.7%), cancer (-3.2%), and controlled diabetes (-2.4%). No other risk factor has an effect of greater than $\pm 2.5\%$ on India's relative mortality. The combined effect of health conditions leads to 9.7% higher mortality in England than in India, reflecting England's higher age-specific prevalence of certain conditions like obesity and cancer, as well as its older age structure that increases population share at ages with worse health. This differential mortality risk does *not* include the direct effect of older age, which is associated with substantial risk (RR_c 4.68 for age 70-80, 11.92 above 80) and magnifies England's mortality disadvantage substantially.

Figure 3
Condition-specific population relative risk in India v. England



Combining the population relative risk from health conditions with the direct effect of demographics on mortality, we modelled the density of deaths across the age distribution (figure 4). In England, 8.8% of expected deaths are below age 60, closely matching the 6.5% observed in England through May 2020 and the 8.6% reported in the OpenSAFELY dataset. In India, 22.0% of modelled deaths are below age 60, which is substantially lower than the 50% observed in case reports. Applying England's age-specific prevalence of health conditions to India's demographic distribution, in order to isolate the effect of health conditions from demographics, results in a distribution nearly identical to the India model. In other words, differences in health conditions between India and England have almost no effect on mortality, indicating that the modelled shift toward younger populations comes from the demographic distribution alone.

Figure 4
Modelled age distribution of COVID-19 mortality



In the appendix, we test sensitivity to uncertainty in prevalences and HRs (appendix pp 6–8). The latter estimates cover alternate hazard ratios estimated from other studies.^{18,26} We also test sensitivity to alternate assumptions about covariance of health conditions (appendix p 9). In all cases, we find that the population relative risk from health conditions in England is greater than in India, and that accounting for health conditions cannot explain any of the higher incidence of mortality among the young in India relative to England.

DISCUSSION

We used the best publicly available data on population health to examine the extent to which demographics and pre-existing health conditions known to increase COVID-19 mortality can account for the disproportionately high share of COVID-19 deaths in younger populations observed in India relative to England. We show that differences in population health do not significantly shift the relative age distribution of disease severity and slightly *lower* aggregate mortality in India relative to England. Higher

prevalence of diabetes and respiratory illness raise mortality risk relative to England, but these effects are offset by lower rates of obesity, heart disease, and cancer. While the Indian age distribution substantially shifts expected mortality toward the young, it explains less than a third of the difference in the share of deaths under 60 compared with England.

Epidemiologic models have typically assumed that comorbidities will exacerbate the mortality of COVID-19 in India and other poorer countries relative to rich countries. We found that comorbidities identified as key risk factors in rich countries do not increase expected mortality in India relative to England, in aggregate or among the young. This suggests that understanding the other factors that may explain the differential mortality among the young observed in lower income contexts, such as different patterns of infection, under-resourced health systems, or comorbidities unique to LMICs, should be a priority for further research.

This study improves upon prior work by examining the extent to which comorbidities can explain the younger incidence of COVID-19 mortality in LMICs, by estimating mortality effects of specific comorbidities, and by calibrating a model with a comprehensive set of comorbidity hazard ratios drawn from a large-sample multivariate analysis of COVID-19 mortality. Models calibrated with bivariate hazard ratios or raw prevalences of comorbidities among severe cases are likely to overstate the effect of pre-existing health conditions because of the significant increase in all comorbidities with age alongside the direct effect of age on COVID-19 mortality.

The key limitation of this study is that there are virtually no data on the COVID-19 mortality risks associated with health conditions that are more common in LMICs than in high income countries, such as protein calorie malnutrition, micronutrient deficiency, and HIV/AIDS.⁴ If these conditions make individuals more susceptible to severe infections, then population health may indeed exacerbate the severity of COVID-19 in LMICs. Understanding the extent to which health conditions endemic to poor countries affect COVID-19 severity is an urgent priority, particularly as policy responses increasingly focus on identifying and isolating high risk individuals.²⁷

Our analysis is also constrained by the limited and changing evidence on risk factors for COVID-19 severity. Based on the availability of existing measures, our model assumed that health condition relative risks are age-invariant. However, data from New York's epidemiological surveillance system suggest that hypertension and diabetes may contribute more to mortality at younger ages,²⁸ which would exacerbate the burden of illness among the young in LMICs. Further, if illness severity and the quality of prior medical management of pre-existing health conditions change mortality risk for the same diagnosis across contexts, applying HRs from England may understate mortality risk in India. Finally, hazard ratios which are not conditioned on infection may reflect infection risk in addition to disease severity risk and thus may not translate directly to the Indian context.

Recognizing these limitations, we have posted our analysis on an open web platform, allowing estimates to be calibrated with different risk factors, hazard ratios, and data from other countries, as more research on the virus emerges.

Data Sharing Statement

All aggregate data and code will be made available in a public repository hosted at github.com and a web platform will replicate the analysis and allow use of alternate assumptions or data. Deidentified Indian health microdata is available from the International Institute for Population Sciences in Mumbai.

Ethics approval

This study was ruled exempt from human subjects review by the Institutional Review Board at Dartmouth College.

Transparency statement

The lead author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported. No important aspects of the study have been omitted and there are no substantial discrepancies from the study as originally planned.

Role of the funding source

Asher and Novosad received funding from Emergent Ventures.

Author Statements

All authors had full access to all the data in the study and shared the final responsibility for the decision to submit for publication. All authors saw and approved the final version of the manuscript.

Author contributions

The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Copyright/license for publication

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, a worldwide licence to the Publishers and its licensees in perpetuity, in all forms, formats and media (whether known now or created in the future), to i) publish, reproduce, distribute, display and store the Contribution, ii) translate the Contribution into other languages, create adaptations, reprints, include within collections and create summaries, extracts and/or, abstracts of the Contribution, iii) create any other derivative work(s) based on the Contribution, iv) to exploit all subsidiary rights in the Contribution, v) the inclusion of electronic links from the Contribution to third party material where-ever it may be located; and, vi) licence any third party to do any or all of the above.

Competing interests

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: Asher and Novosad had financial support from Emerging Ventures for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

References

- 1 Coronavirus disease (COVID-19) Situation Report - 142. World Health Organization, 2020.
- 2 Onder G, Rezza G, Brusaferro S. Case-Fatality Rate and Characteristics of Patients Dying in Relation to COVID-19 in Italy. *JAMA* 2020; **323**: 1775–6.
- 3 Dowd JB, Andriano L, Brazel DM, *et al.* Demographic science aids in understanding the spread and fatality rates of COVID-19. *Proc Natl Acad Sci* 2020; **117**: 9696.
- 4 Walker PGT, Whittaker C, Watson OJ, *et al.* The impact of COVID-19 and strategies for mitigation and suppression in low- and middle-income countries. *Science* 2020; : eabc0035.
- 5 Oliveira MHS de, Wong J, Lippi G, Henry BM. Analysis of clinical and demographic heterogeneity of patients dying from COVID-19 in Brazil versus China and Italy. *Braz J Infect Dis Off Publ Braz Soc Infect Dis* 2020; : S1413-8670(20)30056-8.
- 6 Updates on COVID-19: May 21 2020. Ministry of Health and Family Welfare, Government of India, 2020 <https://pib.gov.in/PressReleaseDetailm.aspx?PRID=1625744> (accessed June 22, 2020).
- 7 McCoy T, Traiano H. In the developing world, the coronavirus is killing far more young people. *Wash. Post.* 2020; published online May 22. https://www.washingtonpost.com/world/the_americas/coronavirus-brazil-killing-young-developing-world/2020/05/22/f76d83e8-99e9-11ea-ad79-eef7cd734641_story.html (accessed June 22, 2020).
- 8 Ioannidis JPA, Axfors C, Contopoulos-Ioannidis DG. Population-level COVID-19 mortality risk for non-elderly individuals overall and for non-elderly individuals without underlying diseases in pandemic epicenters. *medRxiv* 2020; : 2020.04.05.20054361.
- 9 Gilbert M, Pullano G, Pinotti F, *et al.* Preparedness and vulnerability of African countries against importations of COVID-19: a modelling study. *The Lancet* 2020; **395**: 871–7.
- 10 Pearson CAB, van Zandvoort K, Jarvis CI, *et al.* Modelling projections for COVID-19 epidemic in India. LSHTM CMMID COVID-19 Working Group, 2020.
- 11 Richardson S, Hirsch JS, Narasimhan M, *et al.* Presenting Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized With COVID-19 in the New York City Area. *JAMA* 2020; **323**: 2052–9.
- 12 Wu Z, McGoogan JM. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention. *JAMA* 2020; **323**: 1239–42.
- 13 Zhou F, Yu T, Du R, *et al.* Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *The Lancet* 2020; **395**: 1054–62.
- 14 Davies NG, Klepac P, Liu Y, *et al.* Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nat Med* 2020; published online June 16. DOI:10.1038/s41591-020-0962-9.
- 15 van Zandvoort K, Jarvis CI, Pearson C, *et al.* Response strategies for COVID-19 epidemics in African settings: a mathematical modelling study. *medRxiv* 2020; : 2020.04.27.20081711.

- 16 Banerjee A, Pasea L, Harris S, *et al.* Estimating excess 1-year mortality associated with the COVID-19 pandemic according to underlying conditions and age: a population-based cohort study. *The Lancet* 2020; **395**.
- 17 Cabore JW, Karamagi HC, Kipruto H, *et al.* The potential effects of widespread community transmission of SARS-CoV-2 infection in the World Health Organization African Region: a predictive model. *BMJ Glob Health* 2020; **5**: e002647.
- 18 Clark A, Jit M, Warren-Gash C, *et al.* Global, regional, and national estimates of the population at increased risk of severe COVID-19 due to underlying health conditions in 2020: a modelling study. *Lancet Glob Health* 2020; published online June 15. DOI:[https://doi.org/10.1016/S2214-109X\(20\)30264-3](https://doi.org/10.1016/S2214-109X(20)30264-3).
- 19 Williamson E, Walker AJ, Bhaskaran KJ, *et al.* OpenSAFELY: factors associated with COVID-19-related hospital death in the linked electronic health records of 17 million adult NHS patients. *medRxiv* 2020; : 2020.05.06.20092999.
- 20 Petrilli CM, Jones SA, Yang J, *et al.* Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York City: prospective cohort study. *BMJ* 2020; **369**. DOI:10.1136/bmj.m1966.
- 21 Lifestyles Team, NHS Digital. Health Survey for England 2018. 2019 <https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england/2018>.
- 22 World Health Organization and International Diabetes Fund. Definition and diagnosis of diabetes mellitus and intermediate hyperglycemia. World Health Organization, 2006 https://www.who.int/diabetes/publications/Definition%20and%20diagnosis%20of%20diabetes_new.pdf.
- 23 Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2017 (GBD 2017) Results. Seattle, United States: Institute for Health Metrics and Evaluation (IHME), 2018 <http://ghdx.healthdata.org/gbd-results-tool>.
- 24 McLean S, Hoogendoorn M, Hoogenveen RT, *et al.* Projecting the COPD population and costs in England and Scotland: 2011 to 2030. *Sci Rep* 2016; **6**. DOI:10.1038/srep31893.
- 25 Leung JM, Sin DD. Smoking, ACE-2, and COVID-19: Ongoing Controversies. *Eur Respir J* 2020; : 2001759.
- 26 Cummings MJ, Baldwin MR, Abrams D, *et al.* Epidemiology, clinical course, and outcomes of critically ill adults with COVID-19 in New York City: a prospective cohort study. *ET J* 2020; **395**.
- 27 Das J, Sánchez-Páramo C. Smart containment: How low-income countries can tailor their COVID-19 response. World Bank Blogs. 2020; published online May 14. <https://blogs.worldbank.org/voices/smart-containment-how-low-income-countries-can-tailor-their-covid-19-response> (accessed June 20, 2020).
- 28 covid19tracker- Fatalities: Top 10 Comorbidities by Age Group. New York State: Department of Health, State of New York, 2020 <https://covid19tracker.health.ny.gov/views/NYS-COVID19-Tracker/NYSDOHCOVID-19Tracker-Fatalities?%3>.